PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : C12N 15/62, 15/88 — A1

(11) International Publication Number: WO 90/12099

(43) International Publication Date: 18 October 1990 (18.10.90)

(21) International Application Number: PCT/US90/01859

(22) International Filing Date: 5 April 1990 (05.04.90)

(30) Priority data:
335,688 10 April 1989 (10.04.89) US

(71) Applicant: BIOGEN, INC. [US/US]; 14 Cambridge Center, Cambridge, MA 02142 (US).

(72) Inventor: WALLNER, Barbara, P. ; 7 Center Street, Cambridge, MA 02139 (US).

(74) Agents: YANKWICH, Leon, R. et al.; Fish & Neave, 875 Third Avenue, New York, NY 10022-6250 (US).

(81) Designated States: AT (European patent), AU, BE (European patent), CA, CH (European patent), DE (European patent), DK (European patent), ES (European patent), FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: PHOSPHATIDYLINOSITOL LINKING OF PROTEINS

(57) Abstract

DNA sequences derived from a phosphatidylinositol-linked form of lymphocyte function-associated antigen 3 ("LFA-3") are provided which code for a phosphatidylinositol linkage signalling sequence. The linkage signalling sequence may be linked to DNA coding for secretory proteins or polypeptides to obtain phosphatidylinositol-linked chimeric proteins or polypeptides. The chimeric proteins can be used to produce targeted drugs, to form micellular or liposomal drug delivery systems, or to improve the purification or screening of particular cells, proteins or DNA libraries.

DESIGNATIONS OF "DE"

Until further notice, any designation of "DE" in any international application whose international filing date is prior to October 3, 1990, shall have effect in the territory of the Federal Republic of Germany with the exception of the territory of the former German Democratic Republic.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MG	Madagascar
AU	Australia	FI	Finland	ML	Mali
BB	Barbados	FR	France	MR	Mauritania
BE	Belgium	GA	Gabon	MW	Malawi
BF	Burkina Faso	GB	United Kingdom	NL	Netherlands
BG	Bulgaria	HU	Hungary	NO	Norway
BJ	Benin	IT	Italy	RO	Romania
BR	Brazil	JP	Japan	SD	Sudan
CA	Canada	KP	Democratic People's Republic of Korea	SE	Sweden
CF	Central African Republic			SN	Senegal
CG	Congo	KR	Republic of Korea	SU	Soviet Union
CH	Switzerland	LI	Liechtenstein	TD	Chad
CM	Cameroon	LK	Sri Lanka	TG	Togo
DE	Germany, Federal Republic of	LU	Luxembourg	US	United States of America
DK	Denmark	MC	Monaco		

PHOSPHATIDYLINOSITOL LINKING OF PROTEINS

The present invention relates to glycosyl phosphatidylinositol ("PI") linkage structures, which in vivo covalently anchor certain proteins to the surface of the cells in which they are produced. More particularly, the present invention relates to the isolation of DNA sequences coding for the phosphatidylinositol linkage signalling sequence of lymphocyte function-associated antigen 3 ("LFA-3"), and to processes for providing PI linkage structures to extracellular proteins (or the extracellular portion of membrane proteins), thereby giving such proteins a PI-linked form, with resulting advantages such as releasable plasma membrane binding, enhanced purifiability, the capability of micelle formation, etc. The invention also relates to novel chimeric polypeptides having PI linkage structures, hybrid DNA sequences encoding such polypeptides, and to products, methods and compositions made possible by the discoveries detailed herein.

BACKGROUND OF THE INVENTION

Proteins produced by expression of cellular DNA may be divided into three general groups: cytoplasmic proteins, which remain wholly within the cell; extracellular or "soluble" proteins, which are secreted outside the cell; and membrane proteins, which

become attached to the phospholipid bilayer of the cell membrane.

Different membrane proteins play major roles in a wide range of cell functions, such as anchoring cytoskeletal components, mediating cell-to-cell adhesion, transporting molecules into and out of the cell, receiving signals from hormones and other chemical transmitters, and many others.

Membrane proteins are bound to the plasma membrane in a variety of ways. Peripheral or "extrinsic" proteins do not penetrate the phospholipid bilayer but instead interact with the polar head groups on one surface of the bilayer or interact with other membrane proteins anchored directly to the bilayer. Integral or "intrinsic" proteins have a region which interacts with the hydrophobic core of the bilayer. Transmembrane proteins, such as cell surface antigens, cellular receptors, adhesion molecules and transport proteins, traverse the bilayer one or more times, exposing both cytoplasmic and extracellular regions or domains.

A further possibility for attachment to the plasma membrane is via covalent linkages between the lipid bilayer and the protein. In particular, a number of cell surface proteins are anchored to the cell membrane through a C-terminal, covalently attached glycosylated phosphatidylinositol moiety. See, e.g., Low et al., "Covalently Attached Phosphatidylinositol As A Hydrophobic Anchor For Membrane Protein", Trends Biol. Sci., 11, pp. 212-15 (1986). The glycosyl phosphatidylinositol linkage, or "PI" linkage, is believed to be the anchoring structure for more than two dozen specific membrane proteins, found in a wide distribution of species and cell types. See Ferguson et al., "Cell Surface Anchoring Of Proteins Via

Glycosyl-Phosphatidylinositol Structures", Ann. Rev. Biochem., 57, pp. 285-320 (1988) (incorporated herein by reference). The biological functions of these PI-linked proteins are also highly diverse, including surface hydrolases, coat proteins, surface antigens and adhesion molecules.

The C-terminal amino acid residue of the PI-linked protein is attached to the membrane phosphatidylinositol moiety via an ethanolamine-phosphodiester-glycan bridge. Ferguson et al., "Glycosyl-Phosphatidylinositol Moiety That Anchors Trypanosoma brucei Variant Surface Glycoprotein To The Membrane", Science, 239, pp. 753-59 (1988); Low, "Biochemistry Of The Glycosyl-Phosphatidylinositol Membrane Protein Anchors", J. Biochem., 244, pp. 1-13 (1987) (incorporated herein by reference). The mechanism of attachment is thought to involve processing of a C-terminal hydrophobic sequence present on the precursor protein but eliminated prior to the PI anchoring of the mature protein. Ferguson et al., supra, Ann. Rev. Biochem., 57, at 301-04. The processed C-terminal segments are believed to be a signal for phosphatidylinositol attachment. In one study, DNA coding for the 37-amino acid C-terminal sequence of the PI-linked protein, decay accelerating factor (DAF), was fused to the 3' end of DNA coding for a normally secreted protein fragment of glycoprotein D (from herpes simplex virus-1), resulting in a PI-linked fusion protein. See Caras et al., "Signal For Attachment Of A Phospholipid Membrane Anchor In Decay Accelerating Factor", Science, 238, pp. 1280-83 (1987). However, comparison of the C-terminal sequences of many precursors of PI-linked proteins has failed to reveal a consensus PI linkage signalling sequence. See, Low,

supra, Figure 3; Ferguson et al., supra, Ann. Rev. Biochem., 57, Table 3.

Although there has been a considerable amount of investigation of phosphatidylinositol-linked proteins and the phosphatidylinositol linkage, there is a need for further characterization of specific PI linkages and for further investigation into the signal and mechanism for PI attachment to the cellular membrane, as well as the selective release of membrane bound proteins. These needs are addressed by the present invention, relating to the PI linkage structure of a PI-linked form of lymphocyte function-associated antigen 3 ("LFA-3"), and to applications of PI linkage signalling sequences derived from PI-linked LFA-3.

The PI linkage signalling sequences of the present invention, when linked in frame to the 3' end of DNA coding for a secreted polypeptide, or the secreted portion of a polypeptide, result on expression in a PI-linked form of the polypeptide. Using the DNA sequences and methods herein, novel chimeric proteins bearing a C-terminal phosphatidylinositol structure and numerous applications for such products are made possible.

SUMMARY OF THE INVENTION

Accordingly, it is an object of the present invention to provide DNA sequences coding for a phosphatidylinositol linkage signalling sequence.

It is a further object of the present invention to provide DNA sequences coding on expression for novel polypeptides having a C-terminal phosphatidylinositol linkage structure.

It is a further object of the present invention to provide novel PI-linked polypeptides and to provide a means for synthesizing a PI-linked form of

soluble proteins or of the extracellular domains of integral membrane proteins.

It is a further object of the present invention to provide micelles of novel polypeptides bearing intact phosphatidylinositol linkage structures and to provide liposomes having surface characteristics determined by PI-linked polypeptides according to this invention.

It is a further object of the present invention to provide a means for altering the surface protein complement of cell cultures and to provide host cells capable of expressing PI-linked forms of exogenous proteins or polypeptides.

It is a further object of this invention to provide methods for directing cell-to-cell adhesion or targeting specific cells, for screening cells, and for directing the action or increasing the specificity of proteins and therapeutic agents.

These and other objects which will be apparent from the following description are accomplished herein by the discovery of DNA sequences coding for a lymphocyte function-associated antigen 3 phosphatidylinositol linkage signalling sequence, and DNA sequences which code on expression for polypeptides coded for on expression by such DNA sequences.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 depicts a comparison of the DNA insert carried in phage λP24 (and the deduced amino acid sequence), which contains a coding region for PI-linked LFA-3, and the DNA insert carried in phage λHT16 (and the deduced amino acid sequence), which contains a coding region for the transmembrane form of LFA-3.

Figure 2 depicts the DNA coding sequence and the deduced amino acid sequence of PI-linked LFA-3, with underscoring indicating a phosphatidylinositol linkage signalling sequence.

Figure 3 depicts the amino acid sequence of natural CD4 protein, a surface antigen on T-cells. The putative signal sequence and the transmembrane region are indicated by underscoring.

Figure 4 depicts the DNA coding sequence contained in plasmid T4/LFA-3/AD and coding for a fusion protein characterized by an extracellular domain of recombinant CD4 protein and a terminal sequence corresponding to the 28 C-terminal amino acids of PI-linked LFA-3. The deduced amino acid sequence of the chimeric protein is also indicated.

Figure 5 depicts the DNA coding sequence contained in plasmid T4/LFA-3/2 and coding for a chimeric protein characterized by an extracellular domain of recombinant CD4 protein and a terminal sequence believed to represent a PI linkage signalling sequence according to the invention. The deduced amino acid sequence of the chimeric protein is also indicated. The fusion protein resulting from expression of this coding sequence in a transformed host is expected to yield a PI-linked form of CD4.

Figure 6 is a graphic representation of an experiment to show that Clone #11 of CHO cells transfected with the hybrid DNA sequence T4/LFA-3/2 according to the invention produced a PI-linked CD4 protein, releasable from the cell surface by incubation with PIPLC.

DETAILED DESCRIPTION OF THE INVENTION

It has been discovered herein that a portion of the 3' end of cDNA coding for PI-linked LFA-3 can be

isolated and linked, in turn, to the 3' end of a DNA sequence coding for a secreted polypeptide (or a secreted portion of a polypeptide) to confer PI linkage to the polypeptide. The 3' portion selected from the DNA coding for PI-linked LFA-3 which is able to confer PI linkage to a secreted polypeptide is referred to herein as a "phosphatidylinositol linkage signalling sequence." Preferably, the phosphatidylinositol linkage signalling sequences are derived from DNA coding for the PI-linked form of lymphocyte function-associated antigen 3 ("LFA-3") as described herein.

LFA-3 is a surface glycoprotein found on many cells, e.g., erythrocytes, monocytes, granuloctyes, cytotoxic T-lymphocytes, B-lymphoblastoid cells, smooth muscle cells, endothelial cells, fibroblasts. Springer et al., "The Lymphocyte Function-Associated LFA-1, CD2, and LFA-3 Molecules: Cell Adhesion Receptors Of The Immune System", Ann. Rev. Immunol., 5, pp. 223-52 (1987). LFA-3 binds to CD2 (a T-lymphocyte accessory molecule) and is believed to mediate adhesion of T-lymphocytes to target cells. This adhesion is essential to initiation of the T-lymphocyte functional response. Dustin et al., "Purified Lymphocyte Function-Associated Antigen 3 Binds To CD2 And Mediates T Lymphocyte Adhesion", J. Exp. Med., 165, pp. 677-92 (1987).

LFA-3 occurs in two distinct cell surface forms: a transmembrane form and a PI-linked form. Dustin et al., "Anchoring Mechanisms For LFA-3 Cell Adhesion Glycoprotein At Membrane Surface", Nature, 329, pp. 846-48 (1987). Comparison of cDNA sequences encoding the precursor polypeptides of both forms has revealed that the 5' ends of the sequences are identical but that the 3' ends of the sequences differ.

See commonly assigned, copending U.S. patent application Ser. No. 237,309, filed August 26, 1988 (incorporated herein by reference) and Figure 1 herein. The HT16 cDNA coding for the transmembrane form of LFA-3 contains a 3' segment (HT16 Nucleotides 655 through 723, Figure 1) coding for a putative transmembrane region of 23 hydrophobic or uncharged amino acids and a 3' segment (HT16 Nucleotides 724 through 759, Figure 1) coding for a putative cytoplasmic tail of 12 amino acids. The P24 cDNA coding for the PI-linked form of LFA-3 is shorter at the 3' end by thirty nucleotides, with 100% homology with the transmembrane form until the nucleotides coding for the last four amino acids (P24 Nucleotides 718 through 729, Figure 1). Studies with transmembrane LFA-3 have demonstrated that removal of the 12-amino acid cytoplasmic domain converts the protein into a PI-linked protein. Therefore, it is theorized that the functional signal for phosphatidylinositol attachment is contained within the carboxy-terminal sequence of LFA-3.

The phosphatidylinositol linkage signalling sequences according to the present invention may be isolated from DNA coding for PI-linked LFA-3 or synthesized directly using standard techniques.

Preferred phosphatidylinositol linkage signalling sequences according to the present invention are derived from DNA coding for PI-linked LFA-3. Most preferably, the phosphatidylinositol linkage signalling sequence comprises the following:

5'-AGCAATCCATTATTTAATACAACATCATCAATCATTTTGACAACCTGT
ATCCCAAGCAGCGGTCATTCAAGACACAGATATGCACTTATACCCATACCA
TTAGCAGTAATTACAACATGTATTGTGCTGTATATGAATGGTATGTATGCT
TTT-3'.

The ability to attach a phosphatidylinositol structure according to this invention to an otherwise soluble protein or polypeptide leads to a wide variety of uses, some of which are discussed below:

I. Chimeric Polypeptides

The phosphatidylinositol linkage signalling sequences of the present invention may be used to construct hybrid DNA sequences coding for "chimeric" polypeptides having a PI linkage. As used herein, "chimeric" proteins or polypeptides will refer to proteins or polypeptides endowed with a PI structure in accordance with the teachings herein. A "chimeric protein" therefore will typically have a protein component (e.g., corresponding to a soluble protein) and a phosphatidylinositol component. Chimeric proteins are preferably the result of expression of the hybrid DNA molecules described herein. Such hybrid DNA molecules are advantageously prepared by linking the 5' end of a phosphatidylinositol linkage signalling sequence in the correct reading frame to the 3' end of a DNA segment coding for a secretory polypeptide or the secreted portion of a polypeptide. On expression in an appropriate host cell, a chimeric polypeptide or fusion protein is produced having an extracellular domain and a PI linkage anchoring it to the host cell membrane.

Suitable polypeptides for attachment of phosphatidylinositol structures according to this invention may be any polypeptide or fragment thereof which is normally secreted from the cell in which it is produced, or which can be made to be secreted by the attachment of a secretory signal sequence to the N-terminus. Suitable polypeptides will also include the extracellular portions of membrane-bound proteins, for example, where it is desired to produce a PI-linked

form of a protein normally anchored to the cell surface via a transmembrane region and/or a cytoplasmic domain. Suitable types of polypeptides which may be advantageously provided with a phosphatidylinositol structure include cell surface ligands, antibodies to cell surface antigens, hormones, cytotoxins, cell activation ligands, soluble receptors, tumor markers, etc. In a particularly preferred embodiment, DNA encoding a truncated, soluble form of CD4 (also known as "T4"), a surface protein on helper T cells, is fused in frame with an LFA-3 phosphatidylinositol linkage signalling sequence according to this invention, to provide a novel chimeric protein having the CD4 receptor and a C-terminal phosphatidylinositol structure. (See Figure 3 herein.) DNA sequences coding for this fusion protein can be used to transform host cells, and culturing such hosts will result in a population of cells bearing PI-linked CD4 on the host surface membranes.

Since CD4 is now known to be a receptor to which the HIV virus (responsible for acquired immune deficiency syndrome) attaches in the process of infection, the chimeric CD4/phosphatidylinositol protein prepared as described herein offers a potential treatment for HIV infection: the CD4 receptor can be used to target the HIV virus; and through the PI linkage, the CD4 receptor can be expressed on cells which could destroy the virus. In like manner, the CD4 receptor could be fused to a protein toxic to HIV infected cells (or other virus infected cells) and the PI structure employed in the formation of micelles or liposomes (discussed infra) to provide a high concentration drug delivery system.

PI-linked polypeptides are obtained using the hybrid DNA molecules according to the invention by

expressing the hybrid DNA in a host capable of forming the phosphatidylinositol linkage. Any eukaryotic cell capable of forming a phosphatidylinositol linkage may be employed. The selection of host cell will be made on the basis of additional factors, such as stability of cell line, ease of culturing, compatability with other cells or organisms, cytotoxicity toward other cell types, and many other factors, depending on the final application envisioned for the PI-linked polypeptide produced by the transformation, or the application envisioned for the transformed cells themselves. Judicious selection of prospective host organisms will allow those skilled in the art to tailor the features of this invention to their diverse needs, while still relying on the principles described herein.

A wide variety of unicellular host cells and cell lines are suitable, including, for example, mammalian cell cultures, e.g., CHO, R1.1, B-W, and L-M cells, African green monkey cells (including COS1, COS7, BSC1, BSC40, and BMT10 cells), and human cells (including erythrocytes, monocytes, granuloctyes, cytotoxic T-lymphocytes, B-lymphoblastoid cells, smooth muscle cells, endothelial cells, fibroblasts, etc.), and cell lines developed from the foregoing.

The hybrid DNA molecules are expressed by operatively linking them to an expression control sequence in an appropriate expression vector, which in turn is used to transform an appropriate unicellular host. The operative linking of a hybrid DNA sequence of this invention to an expression control sequence includes the provision of a translation start signal in the correct reading frame upstream of the DNA sequence.

Many host-expression vector combinations may be employed in expressing the DNA sequences of this invention. Useful expression vectors may be

constructed from segments of chromosomal, non-chromosomal or synthetic DNA. Suitable vectors include, for example, various known derivatives of SV40 and known bacterial plasmids, e.g., plasmids from E.coli including col E1, pCR1, pBR322, pMB9 and their derivatives, wider host range plasmids, e.g., RP4, phage DNAs (e.g., numerous derivatives of phage λ, such as NM989, and other DNA phages, such as M13 and Filamenteous single stranded DNA phages), yeast plasmids such as the 2μ plasmid or derivatives thereof, and vectors derived from combinations of plasmids and phage DNAs, such as plasmids which have been modified to employ phage DNA or other expression control sequences. For animal cell expression, plasmid BG312, a plasmid containing the major late promoter of adenovirus 2, is preferred.

In addition, any of a wide variety of expression control sequences -- sequences that control the expression of a DNA sequence when operatively linked to it -- may be used in these vectors to express the DNA sequence of this invention. Such useful expression control sequences include, for example, the early and late promoters of SV40 or the adenovirus, the lac system, the trp system, the TAC or TRC system, the major operator and promoter regions of phage λ, the control regions of fd coat protein, the promoter for 3-phosphoglycerate kinase or other glycolytic enzymes, the promoters of acid phosphatase, e.g., Pho5, the promoters of the yeast α-mating factors, and other sequences known to control the expression of genes of prokaryotic or eukaryotic cells or their viruses, and various combinations thereof. For animal cell expression, it is preferred to use an expression control sequence derived from the major late promoter of adenovirus 2.

It should of course be understood that not all vectors and expression control sequences will function equally well to express the DNA sequences of this invention. Neither will all hosts function equally well with the same expression system. However, one of skill in the art may make a selection among these vectors, expression control sequences, and hosts without undue experimentation and without departing from the scope of this invention. For example, in selecting a vector, the host must be considered because the vector must replicate in it. The vector's copy number, the ability to control that copy number, and the expression of any other proteins encoded by the vector, such as antibiotic markers, should also be considered.

In selecting an expression control sequence, a variety of factors should also be considered. These include, for example, the relative strength of the system, its controllability, and its compatibility with the particular DNA sequence, of this invention, particularly as regards potential secondary structures. Unicellular hosts should be selected by consideration of their compatibility with the chosen vector, the toxicity of the product coded on expression by the DNA sequences of this invention to them, their secretion characteristics, their ability to fold proteins correctly, their stability and culturing requirements, and the ease of purification of the products coded on expression by the DNA sequences of this invention.

Within these parameters one of skill in the art may select various vector/expression control system/host combinations that will express the DNA sequences of this invention.

Various agents are known to disrupt the phosphatidylinositol linkage and release the PI-linked

protein from the cell surface. These include, for example, nitrous acid, phosphatidylinositol-specific phospholipases C ("PIPLC"), phosphatidylinositol-specific phospholipases D, and possibly endoglycosidase, phosophodiesterase and proteinase (Low, supra). PIPLC, e.g., from Trypanosoma brucei, has been the agent most extensively studied. PIPLC irreversibly cleaves the PI linkage of PI-linked proteins with a high degree of specificity, creating a soluble protein. Thus, through the use of PIPLC or other agents which cleave the PI linkage, the PI-linked chimeric polypeptides according to the invention may be selectively released from the surface of the host cells in which they are produced.

The invention, therefore, provides a method for obtaining recombinant proteins in soluble form by culturing a host transformed with a hybrid DNA molecule to yield PI-anchored recombinant proteins, then introducing PIPLC or other agent to the culture medium, causing release of the protein in soluble form. Also contemplated are methods for purifying proteins which are difficult to purify by taking advantage of the releasable PI attachment to the cells in which they are produced. For instance, the host cells can be immobilized on a substrate by means of a surface receptor unrelated to the PI-linked protein, and then the protein can be eluted in soluble form using PIPLC or another phosphatidylinositol linkage-disrupting agent as an eluant.

II. Micelles

Chimeric proteins prepared according to the present invention and thus endowed with a terminal phosphatidylinositol structure can be purified with the phosphatidylinositol structure intact by solubilization

of the membrane with detergents. In detergent solution, the chimeric proteins will generally be monomeric and evenly dispersed; however, removal of detergent by dialysis will cause aggregation of the phosphatidylinositol moieties to form micelles or liposomes.

Micelle formation can be induced and controlled by the addition of phospholipids to achieve a specific ratio of protein to phospholipid. The orientation of the chimeric protein components of the micelles can be controlled also, so that the micelles have an outer surface which is predominantly composed of the phospholipid moieties or predominantly composed of the protein moieties of the chimeric proteins. The size of the micelles may also be controlled by varying the detergent employed, the nature of the added phospholipid, or the phospholipid/protein ratio.

Generally, the size of liposomes directly affects the rate at which they are cleared from the bloodstream. For example, smaller liposomes and negatively charged liposomes appear to be more stable and accumulate in the spleen and liver. Thus, the micelles and liposomes prepared from chimeric proteins according to the present invention can be tailored to remain in the bloodstream for a desired period and to be delivered to specific organs. For example, small micelles according to the invention can be formed with an outer surface exhibiting a predominantly negative charge from the PI moiety, and such micelles would be ideal for delivering a particular protein moiety to the spleen or liver.

The chimeric protein micelles can also be structured so that the protein moiety is located on the outside. This type of structure leads to a high concentration of protein activity in the

microenvironment surrounding the micelle. For example, depending on the nature of the protein component of the micellular chimeric protein, the micelle can function as a multivalent ligand or receptor, or as a unitary protein having multiplied activity. Depending on the characteristics of the ligand-receptor interactions, the close proximity of, e.g., multiple ligands on the surface of a micelle according to the invention, could confer a several-fold higher affinity (due to a lower on-off rate) for the ligand than that found for the monomeric ligand protein.

Many therapeutics are only effective when present in higher concentrations and must be injected in large doses to account for dilution and clearance effects. Mixed micelles can be engineered for drug targeting by including a second or third protein (to which the PI moiety has been attached) which specifically localizes the micelles to the target tissue or organ. For example, formation of a mixed micelle having chimeric ligand components and chimeric cytotoxic components, the former having an affinity for the surface receptor of a specific pathogen and the latter being toxic to the pathogen, produces a drug specifically targeted to the pathogen and also produces a therapeutic which localizes the effect of the cytotoxic component. The specific localization will assure a very high concentration of the drug at the site of interaction.

Proteins to be included in micelles can be cell targeting proteins (cell surface ligands, tumor markers, antibodies to tumor surface antigens, etc.), hormones, toxins, cell activation ligands, soluble receptors (e.g., CD4) or any other peptides or protein drugs.

Another approach to a mixed micelle is through the construction of a fusion protein to which the PI structure is attached. The fusion protein will exert plural functions, i.e., will exhibit plural domains corresponding to the protein fragments spliced together to form the fusion protein component. One part of such a fusion protein micelle component may act as a targeting protein (e.g., ligand of a specific receptor) while the other part of the protein may have a specific therapeutic function (e.g., a toxin that lyses specific cells; a cellular activator, a hormone, etc.).

III. Cell Targeting

The chimeric proteins according to this invention can be expressed in specific host cells and used to target that cell to specific tissues or organs.

As outlined above for the therapeutic applications of micelles, human cells with specific functions (killer cells, helper T-cells, etc.) can be engineered to be targeted to a specific site for action. Cytotoxic T-lymphocytes, natural killer cells, macrophages or other types of cells can be transfected with a PI-linked soluble receptor or ligand specific to the target cell (e.g., tumor cell marker transfected into cytotoxic T-lymphocytes). Such targeted cells may advantageously be cultured and specifically activated in vitro.

Most cell surface proteins are transmembrane proteins having a cytoplasmic domain. The cytoplasmic domain often transduces a cellular signal upon interaction with its extracellular ligand. However, by replacing the cytoplasmic domain with the PI linkage it is possible to attach the extracellular portion of the protein without its cytoplasmic signalling sequence and

thus use that molecule strictly as a targeting molecule. Cell targeting in this fashion provides a number of cytolytic cells for site-specific action which normally do not show any specificity.

IV. Screening DNA Libraries

The sequence coding for the PI linkage can be incorporated in an expression vector and used to construct and screen a cDNA library for proteins which are normally secreted.

The phosphatidylinositol linkage signalling sequences of the present invention may be incorporated in expression vectors and used to construct and screen a cDNA library. In this feature, a PI linkage-conferring sequence is included in the expression vector such that the cloned DNA coding for a normally secreted protein is anchored to the host cell surface on expression. The cells harboring the cloned DNA can then be separated, e.g., by the panning technique described by Seed and Aruffo, "Molecular Cloning Of The CD2 Antigen, The T-Cell Erythrocyte Receptor, By A Rapid Immunoselection Procedure", Proc. Natl. Acad. Sci. USA, 84, pp. 3365-69 (1987) (incorporated herein by reference).

The cDNA library is preferably constructed using an animal cell expression vector which has been engineered so that each protein expressed has the PI linkage. The otherwise secreted proteins will thus be attached to the cell surface of the host cell, and cells expressing a particular protein of interest can be isolated using a monoclonal antibody to that protein. The isolation of cDNA for a secreted protein with monoclonal antibodies or a ligand for the protein by currently established techniques is very laborious and inefficient. The proposed technique therefore

represents a great improvement in efficiency and reliability.

The following examples are provided in order to specifically illustrate the features of the present invention. Although the following description relates to a particularly preferred chimeric protein based on the extracellular domain of CD4, the examples are not intended to limit in any way the scope of the invention.

EXAMPLES

The following describes experiments to produce a PI-linked form of CD4 protein by attaching a terminal segment derived from DNA coding for PI-linked LFA-3 to DNA coding for a portion of the extracellular domain of CD4, which is normally (i.e., in wild type human T-cells) a membrane protein. CD4 is a glycoprotein of 458 amino acids, including a 23-amino acid signal sequence (AA_1 through AA_{23}), a 373-amino acid extracellular domain (AA_{24} through AA_{396}), a transmembrane domain (AA_{397} through AA_{418}), and a cytoplasmic tail (AA_{419} through AA_{458}). The amino acid sequence for CD4 is shown in Figure 3. See also, Littman et al., "Corrected CD4 Sequence", Cell, 55, p. 541 (1988). The putative signal sequence, extracellular domain, transmembrane domain and cytoplasmic domain are delineated by underscoring of the signal and transmembrane sequences.

Construction of a CD4/LFA-3 Expression Vector

In the following constructions, all restriction enzymes and other materials were used according to manufacturers' directions, unless otherwise specified.

DNA coding for PI-linked LFA-3 was obtained from plasmid P24, described in the aformentioned commonly assigned, copending United States patent application Ser. No. 237,309. An LB stab of E.coli cells carrying this plasmid was deposited in the In Vitro International, Inc. culture collection in Linthicum, Maryland, USA on July 22, 1988, under accession number IVI-10180.

Plasmid P24 containing the full length cDNA coding for PI-linked LFA-3 (see Figure 2) was digested with restriction enzyme NotI (New England BioLabs, Beverly, MA) and the 849-base pair (bp) fragment isolated by electrophoresis on 1% agar gel. The 849 bp fragment (P24 insert in Figure 1) was digested with restriction enzyme BbvI (New England BioLabs), and the 205 bp fragment (N_{645}-N_{849} in Figure 1) was isolated by electrophoresis, then blunt-ended with DNA polymerase I (Klenow fragment, New England BioLabs) and ligated into the unique BstEII restriction site of expression vector BG391 (accession no. IVI-10151, In Vitro International, Inc.), which incorporates a gene for a recombinant soluble CD4 protein corresponding to amino acids 1-325 in Figure 3. See PCT publication no. WO 89/01940.

The resulting recombinant DNA molecule, designated T4/LFA-3/AD, contained a DNA sequence coding for a polypeptide characterized by the N-terminal 267 amino acids of CD4, followed directly by the 28 C-terminal amino acids of PI-linked LFA-3. The coding sequence and deduced amino acid sequence of T4/LFA-3/AD is set forth in Figure 4.

Plasmid T4/LFA-3/AD was linearized with XmnI and used to cotransfect CHO ($DHFR^-$) cells by the calcium phosphate method with the linearized StuI fragment of plasmid pAdD26, which carries the DHFR gene. See Kaufman and Sharp, Mol. Cell. Biol., 2,

p. 1304 (1982). Transfected cells were grown and selected in alpha(-) Modified Eagles Medium (MEM) (Gibco).

Surface expression of CD4 transfected cells was determined by Fluorescence Activated Cell Sorter (FACS) analysis: 1 x 10^6 transfected cells and control CHO cells were removed from tissue culture dishes by incubation with Hank's BSS buffer, 0.5 EDTA at 4°C for 15 minutes. The detached cells were then pelleted, resuspended in 50 μl of PBN buffer (1 x PBS, 0.5% BSA, 0.1% sodium azide) and incubated with 100 μl of monoclonal antibody OKT4A (Becton Dickinson, Mountainview, CA) on ice for 45 minutes. The cells were pelleted by centrifugation and resuspended twice in PBN buffer. The cells were pelleted again and the cell pellets were resuspended in 100 μl of a 1:50 dilution of FCI (Fluorescein Conjugated Affinity Purified F (ab') 2 Fragment Sheep Anti-Mouse IgG; Cappel Biomedical, Westchester, Pennsylvania, USA) in PBN buffer and incubated on ice for 30 minutes. The cells were layered onto a cushion of 300 μl 100% fetal calf serum and then resuspended in 800 μl of 1 x PBS buffer and the intensity of fluorescence measured by FACS (Becton Dickinson, Mountainview, California, USA).

By FACS analysis, 40 clones were obtained which were positive for surface expression of the chimeric CD4/LFA-3 polypeptide. Four positive clones were cultured in alpha(-) MEM and incubated with a 1:50 dilution of PIPLC (a generous gift of M. Law, Columbia University, New York, New York, USA) in 100 μl MEM, at 37°C for 30 minutes. 1 x 10^6 cells of each clone before and after PIPLC treatment were prepared for FACS analysis as above. The fluorescent intensities of the PIPLC treated clones were identical to the untreated clones, indicating that the CD4 protein was not

released and that expression of the T4/LFA-3/AD plasmid did not result in a chimeric polypeptide having a simple phosphatidylinositol anchor.

A recombinant DNA molecule having a larger segment from the 3' end of LFA-3 cDNA was next constructed. Following the above procedure, a CD4/LFA-3 hybrid DNA sequence was prepared which coded for the N-terminal 266 amino acids of the extracellular domain of CD4 and the C-terminal 51 amino acids of PI-linked LFA-3. The ligation introduced an isoleucine residue. This hybrid DNA molecule was designated T4/LFA-3/2. The coding sequence and deduced amino acid sequence of T4/LFA-3/2 is set forth in Figure 5.

Plasmid T4/LFA-3/2 was linearized with XmnI as above and used to cotransfect 2 x 10^7 CHO ($DHFR^-$) cells with StuI-linearized plasmid pAdD26 by the calcium phosphate method, as above. Transfected cells were grown and selected in alpha(-) MEM.

Sixteen resistant colonies were isolated, expanded and assayed for surface expression of a CD4 protein by FACS as described above. Four clones that were strongly positive for CD4 surface expression, T4/LFA-3/2 #3, #8, #11 and #15, were assayed further for release of surface CD4 after incubation with PIPLC, as described above. Two of the clones showed a decrease in surface CD4 after incubation with PIPLC, indicating that the extracellular portion of the CD4 protein was attached to the cell extracellular portion of the CD4 protein was attached to the cell surface via a phosphatidylinositol linkage (conferred by expression of the 3' phosphatidylinositol linkage signalling sequence excised from the P24 cDNA); two clones did not show decreased surface expression in CD4 after incubation with PIPLC, indicating that the surface CD4 of those particular clones was not PI-linked. Figure 6

shows in the form of graphs the results of the PIPLC incubations on Clone T4-LFA-3-2 #11 (positive for PI-linked CD4 expression) and Clone T4/LFA-3/2 #15 (control, negative for PI-linked expression).

Microorganisms and hybrid DNA molecules according to the invention are exemplified by a culture deposited in the culture collection of In Vitro International, Inc. 611(P) Hammonds Ferry Road, Linthicum, MD, on April 5, 1989. The culture is identified as follows:

Plasmid	Culture	Accession No.
T4/LFA-3/2	E.coli JA221(pT4/LFA-3/2)	IVI-10202

While the products and methods of the present invention have been described herein with reference to particular DNA sequences and polypeptides, a wide range of additional embodiments are contemplated. For instance, although the foregoing description refers to a specific CD4/LFA-3 construct, many similar constructs are apparent from the disclosure, including, for example, longer or shorter portions of the DNA sequence encoding the extracellular domain of CD4, linked to the preferred LFA-3 PI linkage signalling sequence. Also, PI linkage signalling sequences derived from the PI-linked form of LFA-3 which are longer or shorter than the specific sequence of the examples are contemplated, as are nonhomologous DNA sequences which, through the degeneracy of the genetic code, code on expression for the same operative C-terminal peptides as coded for by PI linkage signalling sequences derived directly from PI-linked LFA-3. The present invention of course also contemplates chimeric polypeptides characterized by N-terminal extracellular regions derived from soluble proteins or the extracellular portions of proteins other than CD4. All such additional embodiments and obvious modifications are

within the intended scope of this invention as defined by the appended claims.

CLAIMS:

1. A DNA sequence coding for a lymphocyte function-associated antigen 3 (LFA-3) phosphatidylinositol linkage signalling sequence.

2. A DNA sequence according to claim 1, selected from the group consisting of: 5'-AGCAATCCATTATTTAATACAACATCATCAATCATTTTGACAACCTGT ATCCCAAGCAGCGGTCATTCAAGACACAGATATGCACTTATACCCATACCA TTAGCAGTAATTACAACATGTATTGTGCTGTATATGAATGGTATGTATGCT TTT-3', and DNA sequences which code on expression for a polypeptide coded for on expression by the foregoing DNA sequence.

3. A hybrid DNA sequence comprising (a) a first DNA sequence coding for a secreted protein, a portion of a secreted protein or an extracellular region of a membrane protein, and, downstream of said first DNA sequence, (b) a second DNA sequence coding for a lymphocyte function-associated antigen 3 phosphatidylinositol linkage signalling sequence operatively linked to said first DNA sequence such that expression of the hybrid DNA sequence results in a phosphatidylinositol-linked polypeptide.

4. A hybrid DNA sequence according to claim 3, wherein the second DNA sequence is selected from the group consisting of: 5'-AGCAATCCATTATTTAATACAACATCATCAATCATTTTGACAACCTGT ATCCCAAGCAGCGGTCATTCAAGACACAGATATGCACTTATACCCATACCA TTAGCAGTAATTACAACATGTATTGTGCTGTATATGAATGGTATGTATGCT TTT-3', and DNA sequences which code on expression for a polypeptide coded for on expression by the foregoing DNA sequence.

5. A hybrid DNA sequence according to claim 3, wherein the first DNA sequence codes for a protein or portion of a protein selected from the group consisting of cell surface ligands, antibodies to cell surface antigens, hormones, cytotoxins, cell activation ligands, surface proteases, soluble receptors, and secreted or extrinsic proteins.

6. A hybrid DNA sequence according to claim 5, wherein the first DNA sequence codes for an extracellular domain of CD4.

7. A recombinant DNA molecule comprising (a) a first DNA sequence coding for a secreted protein, a portion of a secreted protein or an extracellular region of a membrane protein, (b) a second DNA sequence downstream of said first DNA sequence coding for a lymphocyte function-associated antigen 3 phosphatidylinositol linkage signalling sequence and operatively linked to said first DNA sequence such that expression of the first and second DNA sequences results in a phosphatidylinositol-linked polypeptide, and (c) an expression control sequence operatively linked to the first DNA sequence.

8. A recombinant DNA molecule according to claim 7, wherein the second DNA sequence is selected from the group consisting of:
5'-AGCAATCCATTATTTAATACAACATCATCAATCATTTTGACAACCTGT
ATCCCAAGCAGCGGTCATTCAAGACACAGATATGCACTTATACCCATACCA
TTAGCAGTAATTACAACATGTATTGTGCTGTATATGAATGGTATGTATGCT
TTT-3', and DNA sequences which code on expression for a polypeptide coded for on expression by the foregoing DNA sequence.

9. A recombinant DNA molecule according to claim 7, wherein the first DNA sequence codes for a protein or portion of a protein selected from the group consisting of cell surface ligands, antibodies to cell surface antigens, hormones, cytotoxins, cell activation ligands, soluble receptors, surface proteases, and secreted or extrinsic proteins.

10. A recombinant DNA molecule according to claim 7, wherein the first DNA sequence codes for an extracellular domain of CD4.

11. A recombinant DNA molecule according to claim 7, wherein the expression control sequence is selected from the group consisting of early and late promoters of SV40 or adenovirus.

12. A eukaryotic host cell culture transfected with a recombinant DNA molecule comprising (a) a first DNA sequence coding for a secreted protein, a portion of a secreted protein or an extracellular region of a membrane protein, (b) a second DNA sequence downstream of said first DNA sequence coding for a lymphocyte function-associated antigen 3 phosphatidylinositol linkage signaling sequence and operatively linked to said first DNA sequence such that expression of the first and second DNA sequences results in a phosphatidylinositol-linked polypeptide, and (c) an expression control sequence operatively linked to the first DNA sequence.

13. A eukaryotic host cell culture according to claim 12, wherein the second DNA sequence is selected from the group consisting of:

5'-AGCAATCCATTATTTAATACAACATCATCAATCATTTTGACAACCTGT ATCCCAAGCAGCGGTCATTCAAGACACAGATATGCACTTATACCCATACCA TTAGCAGTAATTACAACATGTATTGTGCTGTATATGAATGGTATGTATGCT TTT-3', and DNA sequences which code on expression for a polypeptide coded for on expression by the foregoing DNA sequence.

14. A eukaryotic host cell culture according to claim 12, wherein the first DNA sequence codes for a protein or portion of a protein selected from the group consisting of cell surface ligands, antibodies to cell surface antigens, hormones, cytotoxins, cell activation ligands, soluble receptors, surface proteases, and secreted or extrinsic proteins.

15. A eukaryotic host cell culture according to claim 12, wherein the first DNA sequence codes for an extracellular domain of CD4.

16. A eukaryotic host cell culture according to claim 12, wherein the expression control sequence is selected from the group consisting of early and late promoters of SV40 or adenovirus.

17. A eukaryotic host cell culture according to claim 12, wherein the host cell is selected from the group consisting of CHO cells, R1.1 cells, COS cells, killer cells, helper T-cells, cytotoxic T-lymphocytes and macrophages.

18. A polypeptide bearing a terminal phosphatidylinositol structure encoded by a hybrid DNA sequence according to any one of claims 3-6.

19. A micelle consisting essentially of a multiplicity of polypeptides according to claim 18.

20. A micelle according to claim 19, wherein the phosphatidylinositol structures are predominantly oriented on the outer surface of the micelle.

21. A micelle according to claim 19, wherein the phosphatidylinositol structures are predominantly oriented on the inner surface of the micelle.

22. A method for producing a soluble form of a membrane protein comprising the steps of:

(1) operatively linking to a DNA sequence coding for an extracellular portion of a membrane protein a DNA sequence coding for a lymphocyte function-associated antigen 3 phosphatidylinositol linkage signalling sequence, to form a hybrid DNA sequence;

(2) culturing a eukaryotic host cell transfected with an expression vector containing the hybrid DNA sequence, to produce extracellular polypeptides anchored to the cell membranes of the host cells by phosphatidylinositol linkages;

(3) treating the cell culture with a reagent capable of disrupting phosphatidylinositol linkages; and

(4) collecting the soluble form of the polypeptide from the extracellular medium.

23. A method according to claim 22, wherein the phosphatidylinositol linkage signalling sequence is selected from the group consisting of:

5'-AGCAATCCATTATTTAATACAACATCATCAATCATTTTGACAACCTGT
ATCCCAAGCAGCGGTCATTCAAGACACAGATATGCACTTATACCCATACCA

TTAGCAGTAATTACAACATGTATTGTGCTGTATATGAATGGTATGTATGCT TTT-3', and DNA sequences which code on expression for a polypeptide coded for on expression by any of the foregoing DNA sequences.

24. A method according to claim 22, wherein the reagent is selected from nitrous acid, phosphatidylinositol-specific phospholipase C or phosphatidylinositol-specific pospholipase D.

25. A method according to claim 24, wherein the reagent is phosphatidylinositol-specific phospholipase C (PIPLC).

26. A method for producing target-directed host cells comprising the steps of:

(1) isolating a DNA sequence coding for a ligand of a surface receptor on a target cell;

(2) preparing a hybrid DNA sequence comprising the isolated DNA sequence and a DNA sequence coding for a lymphocyte function-associated antigen 3 phosphatidylinositol linkage signalling sequence operatively linked downstream of the isolated DNA sequence; and

(3) transfecting a host cell with an expression vector containing the hybrid DNA sequence, to obtain host cells bearing surface ligands to the target cell receptors.

27. A method for screening a DNA expression library for the DNA of a secreted protein, comprising the steps of:

(1) preparing a series of expression vectors, each vector including an expression control sequence, a DNA sequence coding for a lymphocyte

function-associated antigen 3 phosphatidylinositol linkage signalling sequence, and a DNA isolate from a DNA library, said expression control sequence, DNA sequence and DNA isolate being oriented such that said DNA isolate is operatively linked to the expression control sequence and said DNA sequence is oriented so as to signal on expression phosphatidylinositol linking of any secreted polypeptide encoded by the DNA isolate;

(2) transfecting eukaryotic host cells with the expression vectors prepared according to step (1) and culturing the transfected cells; and

(3) exposing the transfected host cells to a substrate having one or more bound receptors or ligands for or antibodies against the polypeptide coded for by one or more of said DNA isolates.

28. A method for producing micelles comprising a multiplicity of polypeptides bearing terminal phosphatidylinositol structures, comprising the steps of:

(1) isolating a DNA sequence coding for a secreted protein, a portion of a secreted protein or an extracellular region of a membrane protein;

(2) preparing a hybrid DNA sequence comprising the isolated DNA sequence and a DNA sequence coding for a lymphocyte function-associated antigen 3 phosphatidylinositol linkage signalling sequence operatively linked downstream of the isolated DNA sequence;

(3) transfecting and culturing a host cell with an expression vector containing the hybrid DNA sequence, to obtain host cells bearing a phosphatidylinositol-linked form of the polypeptide encoded by the isolated DNA sequence on the host cell surface;

(4) lysing the transfectant cell culture with detergent, to obtain a dispersion containing monomeric polypeptides with intact phosphatidylinositol structures; and

(5) removing the detergent by dialysis, optionally adding phospholipid, until micellular aggregates of the monomeric polypeptides with intact phosphatidylinositol structures are formed.

COMPARISON OF HT16 AND P24 cDNAs

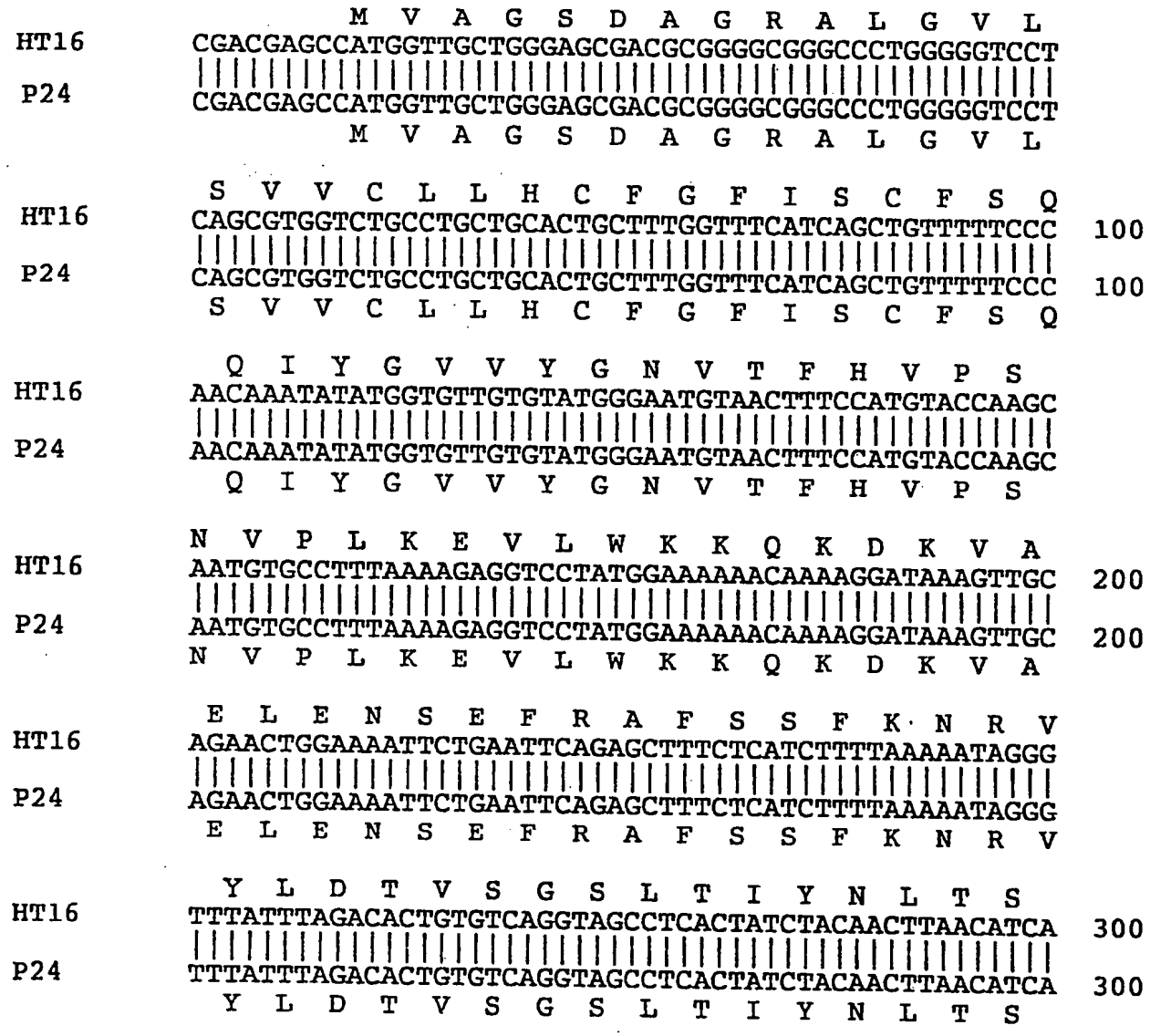


FIG. 1

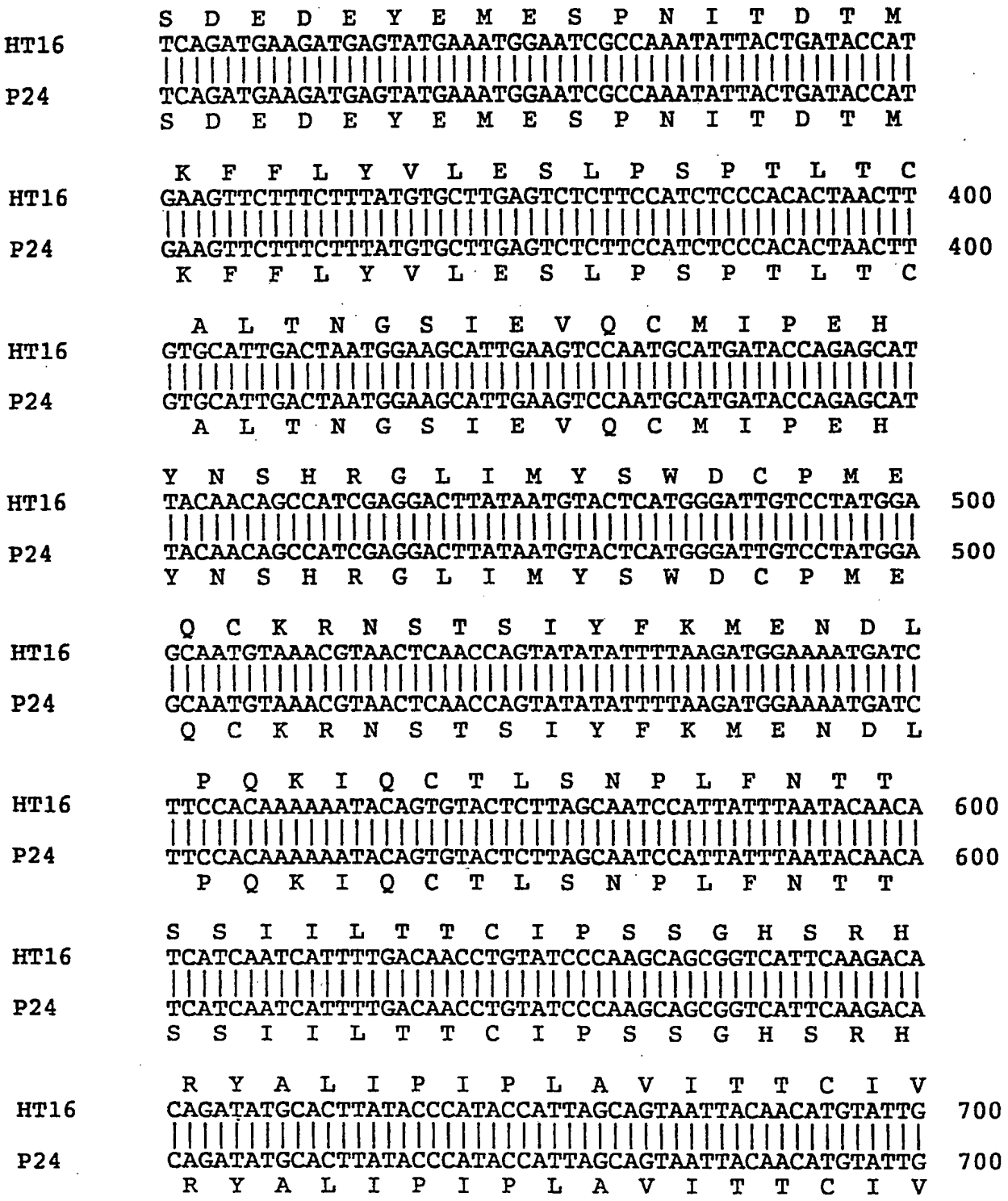

S D E D E Y E M E S P N I T D T M
HT16 TCAGATGAAGATGAGTATGAAATGGAATCGCCAAATATTACTGATACCAT
P24 TCAGATGAAGATGAGTATGAAATGGAATCGCCAAATATTACTGATACCAT
S D E D E Y E M E S P N I T D T M
K F F L Y V L E S L P S P T L T C
HT16 GAAGTTCTTTCTTTATGTGCTTGAGTCTCTTCCATCTCCCACACTAACTT 400
P24 GAAGTTCTTTCTTTATGTGCTTGAGTCTCTTCCATCTCCCACACTAACTT 400
K F F L Y V L E S L P S P T L T C
A L T N G S I E V Q C M I P E H
HT16 GTGCATTGACTAATGGAAGCATTGAAGTCCAATGCATGATACCAGAGCAT
P24 GTGCATTGACTAATGGAAGCATTGAAGTCCAATGCATGATACCAGAGCAT
A L T N G S I E V Q C M I P E H
Y N S H R G L I M Y S W D C P M E
HT16 TACAACAGCCATCGAGGACTTATAATGTACTCATGGGATTGTCCTATGGA 500
P24 TACAACAGCCATCGAGGACTTATAATGTACTCATGGGATTGTCCTATGGA 500
Y N S H R G L I M Y S W D C P M E
Q C K R N S T S I Y F K M E N D L
HT16 GCAATGTAAACGTAACTCAACCAGTATATATTTTAAGATGGAAAATGATC
P24 GCAATGTAAACGTAACTCAACCAGTATATATTTTAAGATGGAAAATGATC
Q C K R N S T S I Y F K M E N D L
P Q K I Q C T L S N P L F N T T
HT16 TTCCACAAAAAATACAGTGTACTCTTAGCAATCCATTATTTAATACAACA 600
P24 TTCCACAAAAAATACAGTGTACTCTTAGCAATCCATTATTTAATACAACA 600
P Q K I Q C T L S N P L F N T T
S S I I L T T C I P S S G H S R H
HT16 TCATCAATCATTTTGACAACCTGTATCCCAAGCAGCGGTCATTCAAGACA
P24 TCATCAATCATTTTGACAACCTGTATCCCAAGCAGCGGTCATTCAAGACA
S S I I L T T C I P S S G H S R H
R Y A L I P I P L A V I T T C I V
HT16 CAGATATGCACTTATACCCATACCATTAGCAGTAATTACAACATGTATTG 700
P24 CAGATATGCACTTATACCCATACCATTAGCAGTAATTACAACATGTATTG 700
R Y A L I P I P L A V I T T C I V

FIG. 1 (cont.)

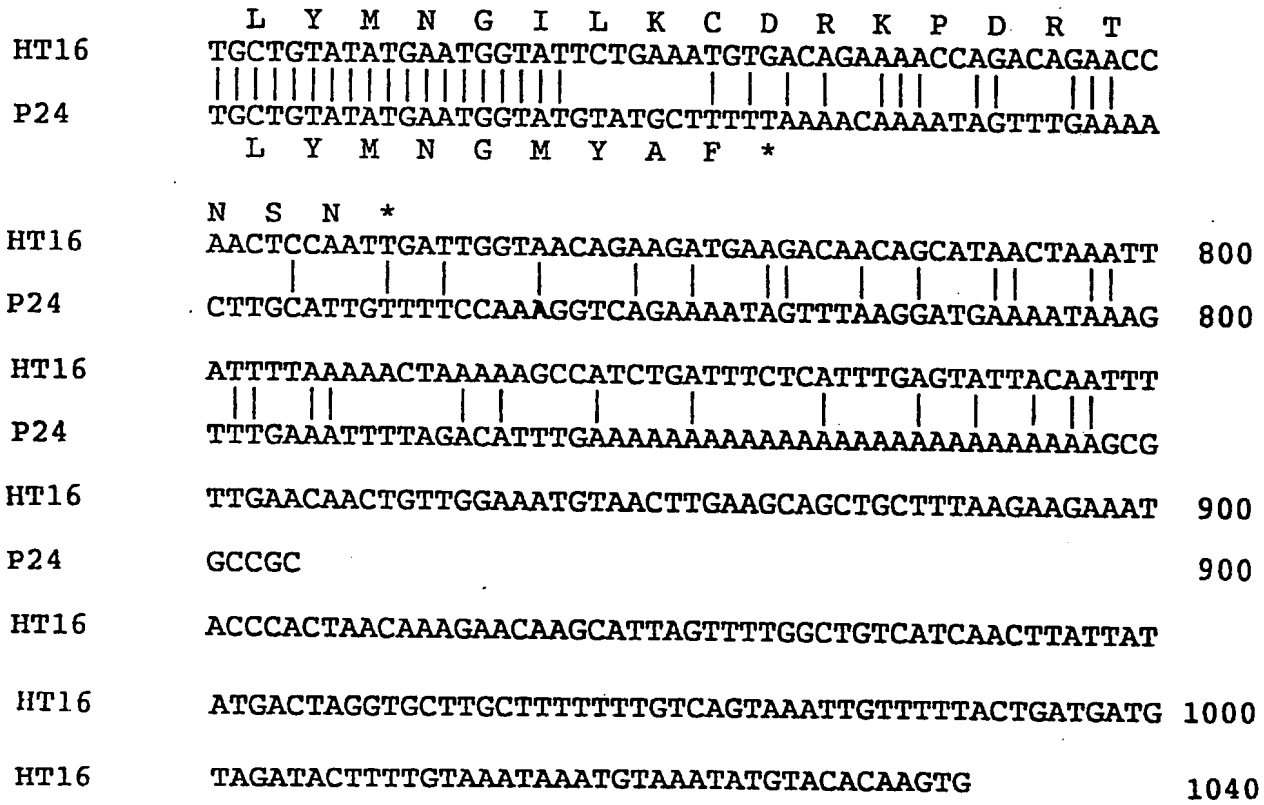

L Y M N G I L K C D R K P D R T
HT16
TGCTGTATATGAATGGTATTCTGAAATGTGACAGAAAACCAGACAGAACC
P24
TGCTGTATATGAATGGTATGTATGCTTTTTAAAACAAAATAGTTTGAAAA
L Y M N G M Y A F *
N S N *
HT16
AACTCCAATTGATTGGTAACAGAAGATGAAGACAACAGCATAACTAAATT 800
P24
CTTGCATTGTTTTCCAAAGGTCAGAAAATAGTTTAAGGATGAAAATAAAG 800
HT16
ATTTTAAAAACTAAAAAGCCATCTGATTTCTCATTTGAGTATTACAATTT
P24
TTTGAAATTTTAGACATTTGAAAAAAAAAAAAAAAAAAAAAAAAAGCG
HT16
TTGAACAACTGTTGGAAATGTAACTTGAAGCAGCTGCTTTAAGAAGAAAT 900
P24
GCCGC 900
HT16
ACCCACTAACAAAGAACAAGCATTAGTTTTGGCTGTCATCAACTTATTAT
HT16
ATGACTAGGTGCTTGCTTTTTTTGTCAGTAAATTGTTTTTACTGATGATG 1000
HT16
TAGATACTTTTGTAAATAAATGTAAATATGTACACAAGTG 1040

FIG. 1 (cont.)

DNA AND AMINO ACID SEQUENCE OF P24

```
1   GCGGCCGCCGACGAGCCATGGTTGCTGGGAGCGACGCGGGGCGGGCCCTGGGGGTCCTCAGCGTGGTCTGCCTGCTGCACTGCTTTGGTTTCATCAGCTG   100
                    MetValAlaGlySerAspAlaGlyArgAlaLeuGlyValLeuSerValValCysLeuLeuHisCysPheGlyPheIleSerCy

101 TTTTTCCCAACAAATATATGGTGTTGTGTATGGGAATGTAACTTTCCATGTACCAAGCAATGTGCCTTTAAAAGAGGTCCTATGGAAAAAACAAAAGGAT   200
    sPheSerGlnGlnIleTyrGlyValValTyrGlyAsnValThrPheHisValProSerAsnValProLeuLysGluValLeuTrpLysLysGlnLysAsp

201 AAAGTTGCAGAACTGGAAAATTCTGAATTCAGAGCTTTCTCATCTTTTAAAAATAGGGTTTATTTAGACACTGTGTCAGGTAGCCTCACTATCTACAACT   300
    LysValAlaGluLeuGluAsnSerGluPheArgAlaPheSerSerPheLysAsnArgValTyrLeuAspThrValSerGlySerLeuThrIleTyrAsnL

301 TAACATCATCAGATGAAGATGAGTATGAAATGGAATCGCCAAATATTACTGATACCATGAAGTTCTTTCTTTATGTGCTTGAGTCTCTTCCATCTCCCAC   400
    euThrSerSerAspGluAspGluTyrGluMetGluSerProAsnIleThrAspThrMetLysPhePheLeuTyrValLeuGluSerLeuProSerProTh

401 ACTAACTTGTGCATTGACTAATGGAAGCATTGAAGTCCAATGCATGATACCAGAGCATTACAACAGCCATCGAGGACTTATAATGTACTCATGGGATTGT   500
    rLeuThrCysAlaLeuThrAsnGlySerIleGluValGlnCysMetIleProGluHisTyrAsnSerHisArgGlyLeuIleMetTyrSerTrpAspCys

501 CCTATGGAGCAATGTAAACGTAACTCAACCAGTATATATTTTAAGATGGAAAATGATCTTCCACAAAAAATACAGTGTACTCTTAGCAATCCATTATTTA   600
    ProMetGluGlnCysLysArgAsnSerThrSerIleTyrPheLysMetGluAsnAspLeuProGlnLysIleGlnCysThrLeuSerAsnProLeuPheA

601 ATACAACATCATCAATCATTTTGACAACCTGTATCCCAAGCAGCGGTCATTCAAGACACAGATATGCACTTATACCCATACCATTAGCAGTAATTACAAC   700
    snThrThrSerSerIleIleLeuThrThrCysIleProSerSerGlyHisSerArgHisArgTyrAlaLeuIleProIleProLeuAlaValIleThrTh

701 ATGTATTGTGCTGTATATGAATGGTATGTATGCTTTTTAAAACAAAATAGTTTGAAAACTTGCATTGTTTTCCAAAGGTCAGAAAATAGTTTAAGGATGA   800
    rCysIleValLeuTyrMetAsnGlyMetTyrAlaPhe

801 AAATAAAGTTTGAAATTTTAGACATTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGGCCGC  863
```

FIG. 2

AMINO ACID SEQUENCE OF FULL LENGTH HUMAN CD4

```
   1  MetAsnArgGlyValProPheArgHisLeuLeuLeuValLeuGlnLeuAlaLeuLeuProAlaAlaThrGlnGlyLysLysValValLeuGlyLysLysG
 101  lyAspThrValGluLeuThrCysThrAlaSerGlnLysLysSerIleGlnPheHisTrpLysAsnSerAsnGlnIleLysIleLeuGlyAsnGlnGlySe
 201  rPheLeuThrLysGlyProSerLysLeuAsnAspArgAlaAspSerArgArgSerLeuTrpAspGlnGlyAsnPheProLeuIleIleLysAsnLeuLys
 301  IleGluAspSerAspThrTyrIleCysGluValGluAspGlnLysGluGluValGlnLeuLeuValPheGlyLeuThrAlaAsnSerAspThrHisLeuL
 401  euGlnGlyGlnSerLeuThrLeuThrLeuGluSerProProGlySerSerProSerValGlnCysArgSerProArgGlyLysAsnIleGlnGlyGlyLy
 501  sThrLeuSerValSerGlnLeuGluLeuGlnAspSerGlyThrTrpThrCysThrValLeuGlnAsnGlnLysLysValGluPheLysIleAspIleVal
 601  ValLeuAlaPheGlnLysAlaSerSerIleValTyrLysLysGluGlyGluGlnValGluPheSerPheProLeuAlaPheThrValGluLysLeuThrG
 701  lySerGlyGluLeuTrpTrpGlnAlaGluArgAlaSerSerSerLysSerTrpIleThrPheAspLeuLysAsnLysGluValSerValLysArgValTh
 801  rGlnAspProLysLeuGlnMetGlyLysLysLeuProLeuHisLeuThrLeuProGlnAlaLeuProGlnTyrAlaGlySerGlyAsnLeuThrLeuAla
 901  LeuGluAlaLysThrGlyLysLeuHisGlnGluValAsnLeuValValMetArgAlaThrGlnLeuGlnLysAsnLeuThrCysGluValTrpGlyProT
1001  hrSerProLysLeuMetLeuSerLeuLysLeuGluAsnLysGluAlaLysValSerLysArgGluLysAlaValTrpValLeuAsnProGluAlaGlyMe
2001  tTrpGlnCysLeuLeuSerAspSerGlyGlnValLeuLeuGluSerAsnIleLysValLeuProThrTrpSerThrProValGlnProMetAlaLeuIle
3001  ValLeuGlyGlyValAlaGlyLeuLeuLeuPheIleGlyLeuGlyIlePhePheCysValArgCysArgHisArgArgArgGlnAlaGluArgMetSerG
4001  lnIleLysArgLeuLeuSerGluLysLysThrCysGlnCysProHisArgPheGlnLysThrCysSerProIle
```

FIG. 3

T4LFA-3AD amino acid and DNA sequence

```
1       ATGAACCGGGGAGTCCCTTTTAGGCACTTGCTTCTGGTGCTGCAACTGGC
        MetAsnArgGlyValProPheArgHisLeuLeuLeuValLeuGlnLeuAl

        GCTCCTCCCAGCAGCCACTCAGGGAAAGAAAGTGGTGCTGGGCAAAAAAG   100
        aLeuLeuProAlaAlaThrGlnGlyLysLysValValLeuGlyLysLysG

101     GGGATACAGTGGAACTGACCTGTACAGCTTCCCAGAAGAAGAGCATACAA
        lyAspThrValGluLeuThrCysThrAlaSerGlnLysLysSerIleGln

        TTCCACTGGAAAAACTCCAACCAGATAAAGATTCTGGGAAATCAGGGCTC   200
        PheHisTrpLysAsnSerAsnGlnIleLysIleLeuGlyAsnGlnGlySe

201     CTTCTTAACTAAAGGTCCATCCAAGCTGAATGATCGCGCTGACTCAAGAA
        rPheLeuThrLysGlyProSerLysLeuAsnAspArgAlaAspSerArgA

        GAAGCTTGTGGGACCAAGGAAACTTTCCCCTGATCATCAAGAATCTTAAG   300
        rgSerLeuTrpAspGlnGlyAsnPheProLeuIleIleLysAsnLeuLys

301     ATAGAAGACTCAGATACTTACATCTGTGAAGTGGAGGACCAGAAGGAGGA
        IleGluAspSerAspThrTyrIleCysGluValGluAspGlnLysGluGl

        GGTGCAATTGCTAGTGTTCGGATTGACTGCCAACTCTGACACCCACCTGC   400
        uValGlnLeuLeuValPheGlyLeuThrAlaAsnSerAspThrHisLeuL

401     TTCAGGGGCAGAGCCTGACCCTGACCTTGGAGAGCCCCCCTGGTAGTAGC
        euGlnGlyGlnSerLeuThrLeuThrLeuGluSerProProGlySerSer

        CCCTCAGTGCAATGTAGGAGTCCAAGGGGTAAAAACATACAGGGGGGGAA   500
        ProSerValGlnCysArgSerProArgGlyLysAsnIleGlnGlyGlyLy

501     GACCCTCTCCGTGTCTCAGCTGGAGCTCCAGGATAGTGGCACCTGGACAT
        sThrLeuSerValSerGlnLeuGluLeuGlnAspSerGlyThrTrpThrC

        GCACTGTCTTGCAGAACCAGAAGAAGGTGGAGTTCAAAATAGACATCGTG   600
        ysThrValLeuGlnAsnGlnLysLysValGluPheLysIleAspIleVal

601     GTGCTAGCTTTCCAGAAGGCCTCCAGCATAGTCTATAAGAAAGAGGGGGA
        ValLeuAlaPheGlnLysAlaSerSerIleValTyrLysLysGluGlyGl
```

FIG. 4

```
                ACAGGTGGAGTTCTCCTTCCCACTCGCCTTTACAGTTGAAAAGCTGACGG  700
                uGlnValGluPheSerPheProLeuAlaPheThrValGluLysLeuThrG

                .    .    .    .    .    .    .    .    .    .
701             GCAGTGGCGAGCTGTGGTGGCAGGCGGAGAGGGCTTCCTCCTCCAAGTCT
                lySerGlyGlyLeuTrpTrpGlnAlaGluArgAlaSerSerSerLysSer

                .    .    .    .    .    .    .    .    .    .
                TGGATCACCTTTGACCTGAAGAACAAGGAAGTGTCTGTAAAACGGGTTAC  800
                TrpIleThrPheAspLeuLysAsnLysGluValSerValLysArgValTh

                .    .    .    .    .    .    .    .    .    .
801             AAGACACAGATATGCACTTATACCCATACCATTAGCAGTAATTACAACAT
                rArgHisArgTyrAlaLeuIleProIleProLeuAlaValIleThrThrC

                .    .    .    .    .    .    .    .    .    .
                GTATTGTGCTGTATATGAATGGTATGTATGCTTTTTAAAACAAAATAGTT  900
                ysIleValLeuTyrMetAsnGlyMetTyrAlaPhe

                .    .    .    .    .    .    .    .    .    .
901             TGAAAACTTGCATTGTTTTCCAAAGGTCAGAAAATAGTTTAAGGATGAAA

                .    .    .    .    .    .    .    .    .    .
                ATAAAGTTTGAAATTTTAGACATTTGAAAAAAAAAAAAAAAAAAAAAAA 1000

1001            AAAGCGGCC 1009
```

FIG. 4 (cont.)

DNA AND AMINO ACID SEQUENCE OF T4/LFA-3/2

```
1   ATGAACCGGGGAGTCCCTTTTAGGCACTTGCTTCTGGTGCTGCAACTGGCGCTCCTCCCAGCAGCCACTCAGGGAAAGAAAGTGGTGCTGGGCAAAAAAG   100
    MetAsnArgGlyValProPheArgHisLeuLeuLeuValLeuGlnLeuAlaLeuLeuProAlaAlaThrGlnGlyLysLysValValLeuGlyLysLysG

101 GGGATACAGTGGAACTGACCTGTACAGCTTCCCAGAAGAAGAGCATACAATTCCACTGGAAAAACTCCAACCAGATAAAGATTCTGGGAAATCAGGGCTC   200
    lyAspThrValGluLeuThrCysThrAlaSerGlnLysLysSerIleGlnPheHisTrpLysAsnSerAsnGlnIleLysIleLeuGlyAsnGlnGlySe

201 CTTCTTAACTAAAGGTCCATCCAAGCTGAATGATCGCGCTGACTCAAGAAGAAGCTTGTGGGACCAAGGAAACTTTCCCCTGATCATCAAGAATCTTAAG   300
    rPheLeuThrLysGlyProSerLysLeuAsnAspArgAlaAspSerArgArgSerLeuTrpAspGlnGlyAsnPheProLeuIleIleLysAsnLeuLys

301 ATAGAAGACTCAGATACTTACATCTGTGAAGTGGAGGACCAGAAGGAGGAGGTGCAATTGCTAGTGTTCGGATTGACTGCCAACTCTGACACCCACCTGC   400
    IleGluAspSerAspThrTyrIleCysGluValGluAspGlnLysGluGluValGlnLeuLeuValPheGlyLeuThrAlaAsnSerAspThrHisLeuL

401 TTCAGGGGCAGAGCCTGACCCTGACCTTGGAGAGCCCCCCTGGTAGTAGCCCCTCAGTGCAATGTAGGAGTCCAAGGGGTAAAAACATACAGGGGGGGAA   500
    euGlnGlyGlnSerLeuThrLeuThrLeuGluSerProProGlySerSerProSerValGlnCysArgSerProArgGlyLysAsnIleGlnGlyGlyLy

501 GACCCTCTCCGTGTCTCAGCTGGAGCTCCAGGATAGTGGCACCTGGACATGCACTGTCTTGCAGAACCAGAAGAAGGTGGAGTTCAAAATAGACATCGTG   600
    sThrLeuSerValSerGlnLeuGluLeuGlnAspSerGlyThrTrpThrCysThrValLeuGlnAsnGlnLysLysValGluPheLysIleAspIleVal
```

FIG. 5

601 GTGCTAGCTTTCCAGAAGGCCTCCAGCATAGTCTATAAGAAAGAGGGGGAACAGGTGGAGTTCTCCTTCCCACTCGCCTTTACAGTTGAAAAGCTGACGG 700
ValLeuAlaPheGlnLysAlaSerSerIleValTyrLysLysGluGlyGluGlnValGluPheSerPheProLeuAlaPheThrValGluLysLeuThrG

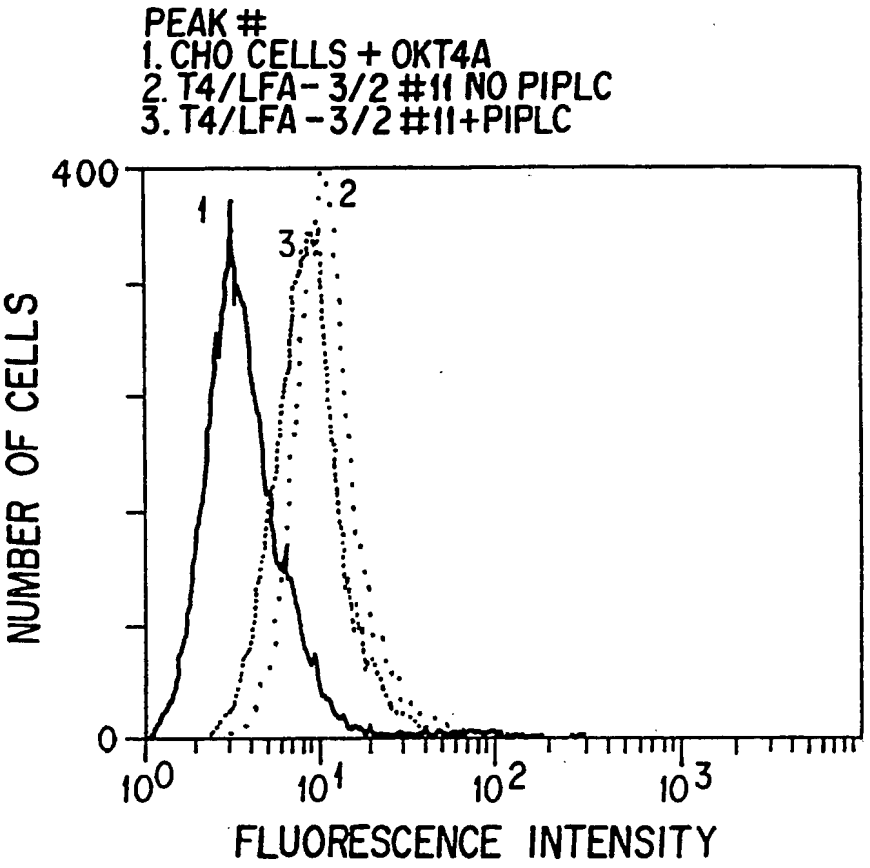
701 GCAGTGGCGAGCTGTGGTGGCAGGCGGAGAGGGCTTCCTCCTCCAAGTCTTGGATCACCTTTGACCTGAAGAACAAGGAAGTGTCTGTAAAACGGGTTAT 800
lySerGlyGluLeuTrpTrpGlnAlaGluArgAlaSerSerSerLysSerTrpIleThrPheAspLeuLysAsnLysGluValSerValLysArgValIl

801 TAGCAATCCATTATTTAATACAACATCATCAATCATTTTGACAACCTGTATCCCAAGCAGCGGTCATTCAAGACACAGATATGCACTTATACCCATACCA 900
eSerAsnProLeuPheAsnThrThrSerSerIleIleLeuThrThrCysIleProSerSerGlyHisSerArgHisArgTyrAlaLeuIleProIlePro

901 TTAGCAGTAATTACAACATGTATTGTGCTGTATATGAATGGTATGTATGCTTTTTAAAACAAAATAGTTTGAAAACTTGCATTGTTTTCCAAAGGTCAGA 1000
LeuAlaValIleThrThrCysIleValLeuTyrMetAsnGlyMetTyrAlaPhe

001 AAATAGTTTAAGGATGAAAATAAAGTTTGAAATTTTAGACATTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGGCC 1078

FIG. 5 (cont)


PEAK #
1. CHO CELLS + OKT4A
2. T4/LFA-3/2 #11 NO PIPLC
3. T4/LFA-3/2 #11+PIPLC
400
0
1
2
3
NUMBER OF CELLS
10^0
10^1
10^2
10^3
FLUORESCENCE INTENSITY

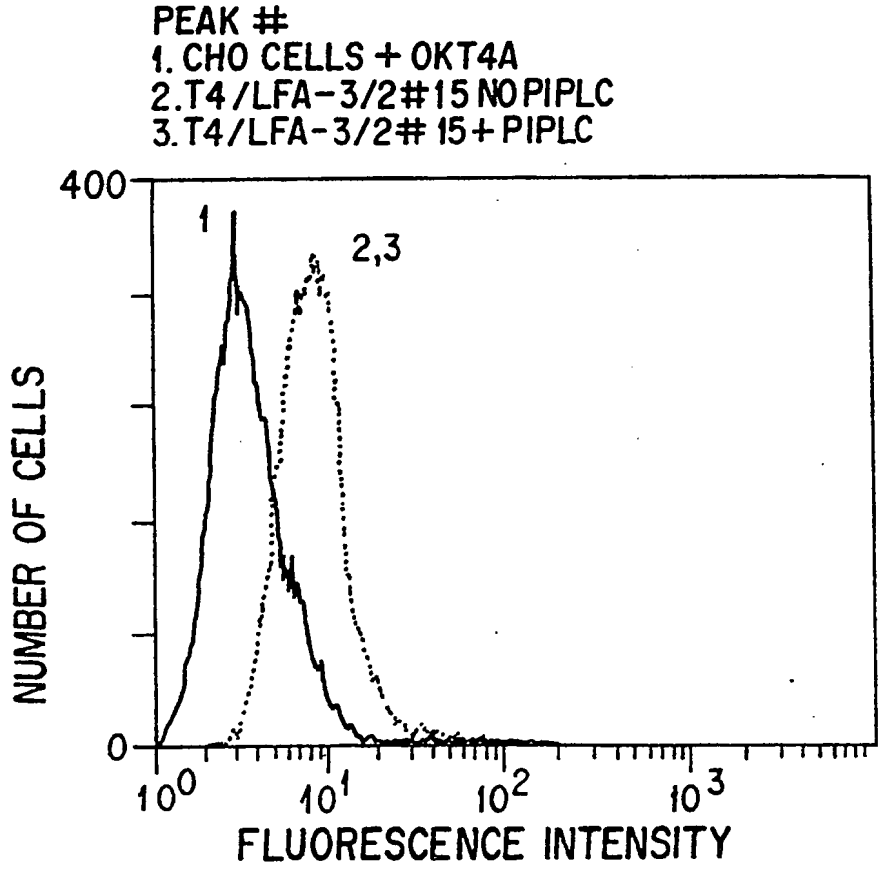

PEAK #
1. CHO CELLS + OKT4A
2. T4/LFA-3/2 #15 NO PIPLC
3. T4/LFA-3/2 # 15 + PIPLC
400
0
1
2,3
NUMBER OF CELLS
10^0
10^1
10^2
10^3
FLUORESCENCE INTENSITY

FIG. 6

MICROORGANISMS

References to deposited microorganisms under PCT Rule 13bis received by the International Bureau on 26 September 1990 (26.09.90).

International Application No: PCT/ US90/01859

MICROORGANISMS

Optional Sheet in connection with the microorganism referred to on page 22, lines 9-10 of the description

A. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet ☒ 1 additional sheet attached

Name of depositary institution

In Vitro International, Inc.

Address of depositary institution (including postal code and country)

611 (P) Hammonds Ferry Road
Linthicum, Maryland 21090
United States of America

Date of deposit	Accession Number
See attached additional sheet.	See attached additional sheet.

B. ADDITIONAL INDICATIONS (leave blank if not applicable). This information is continued on a separate attached sheet ☐

In respect of those designations in which a European patent is sought samples of the deposited microorganisms will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).

C. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)

D. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)

The indications listed below will be submitted to the International Bureau later (Specify the general nature of the indications e.g., "Accession Number of Deposit")

E. ☐ This sheet was received with the international application when filed (to be checked by the receiving Office)

(Authorized Officer)

☐ The date of receipt (from the applicant) by the International Bureau

26 SEPTEMBER 1990 (26.09.90)

J. L. Baron

(Authorized Officer)

Form PCT/RO/134 (January 1981)

Additional Sheet 1 of 1 To Form
PCT/RO/134

Continuation of Box A

IDENTIFICATION OF DEPOSIT

T4/LFA-3/2 : E.coli JA221 (pT4/LFA-3/2)

DATE OF DEPOSIT

5 April 1989

ACCESSION NUMBER

IVI 10202

INTERNATIONAL SEARCH REPORT

International Application No PCT/US90/01859

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [3]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC (5): C12N 15/62; C12N 15/88

U.S.Cl.: 435/172.3; 530/359

II. FIELDS SEARCHED

Minimum Documentation Searched [4]

Classification System	Classification Symbols
U.S.	935/6, 9, 10, 11; 435/172.3, 272

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [6]

Chemical Abstracts Services Online (File CA, 196701988; File BIOSIS Previews 1969-1988). Automated Patent System (File U.S.Pat, 1975-1988).

III. DOCUMENTS CONSIDERED TO BE RELEVANT [14]

Category *	Citation of Document, [16] with indication, where appropriate, of the relevant passages [17]	Relevant to Claim No. [18]
X / Y	Journal of Experimental Medicine, Volume 166, issued October 1987, B. P. Wallner, et al., "Primary Structure of Lymphocyte Function-Associated Antigen 3 (LFA-3) The Ligand of the T Lymphocyte CD2 Glycoprotein," See page 930.	1 / 2-28
Y	Nature, Volume 329, issued 29 October 1987, M.L. Dustin, et al., "Anchoring mechanisms for LFA-3 cell adhesion glycoprotein at membrane surface" See page 846.	1-28
Y	Science, Volume 238, issued 27 November 1987, I. W. Caras, et al., "Signal for Attachment of a Phospholipid Membrane Anchor in Decay Accelerating Factor" See pages 1280-1282.	1-28
X / Y	Nature, Volume 329, issued 29 October 1987, B. Seed, "An LFA-3 cDNA encodes a phospholipid-linked membrane protein homologous to its receptor CD2, See page 840.	1 / 2-28
Y	Proceedings of National Academy of Science USA, Colume 85, issued January 1988, G. L. Wanneck, "Molecular mapping of signals in the Qa-2 antigen (continued on second sheet)	1-28

* Special categories of cited documents: [15]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search [2]	Date of Mailing of this International Search Report [2]
27 June 1990	31 AUG 1990
International Searching Authority [1]	Signature of Authorized Officer [11]
ISA/US	for Laurie A. Scheiner NGUYEN NGOC-HO INTERNATIONAL DIVISION

FURTHER INFORMATION CONTINUED FROM THE SECOND SHEET

required for attachment of the phosphatidylinositol membrane anchor, See page 577.

V. ☐ OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE [1]

This International search report has not been established in respect of certain claims under Article 17(2) (a) for the following reasons:

1. ☐ Claim numbers . . , because they relate to subject matter [1] not required to be searched by this Authority, namely:

2. ☐ Claim numbers . . . . . , because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out [1], specifically:

3. ☐ Claim numbers_____, because they are dependent claims not drafted in accordance with the second and third sentences of PCT Rule 6.4(a).

VI. ☐ OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING [2]

This International Searching Authority found multiple inventions in this international application as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims of the International application.

2. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims of the international application for which fees were paid, specifically claims:

3. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:

4. ☐ As all searchable claims could be searched without effort justifying an additional fee, the International Searching Authority did not invite payment of any additional fee.

Remark on Protest

☐ The additional search fees were accompanied by applicant's protest.

☐ No protest accompanied the payment of additional search fees.